Supplementary Information
Analysis and Refactoring of the A-74528 Biosynthetic Pathway

Jay T. Fitzgerald, Louise K. Charkoudian, Katharine R. Watts, Chaitan Khosla

MATERIALS AND METHODS

General

Deuterated solvents for NMR experiments were purchased from Cambridge Isotope Laboratories (Andover, MA, USA) and all other solvents were purchased from Fisher Scientific (Pittsburgh, PA, USA) at the highest available grade. ^{1}H and ^{13}C NMR spectra of purified polyketide products were recorded on Varian Inova 600 MHz and a Varian Inova 500 MHz (125 MHz for ^{13}C) respectively (Varian, Walnut Creek, CA, USA) in CD_3OD or DMSO-d_6 as indicated. The ^{1}H NMR spectra were referenced to the solvent peak at 3.31 p.p.m. for CD_3OD or 2.50 p.p.m. for DMSO-d_6. LC/MS spectra were obtained on an Agilent 1260 Infinity HPLC with a Agilent 6520 Q-TOF mass spectrometer. The column used was a Phenomenex Gemini®-NX 5 μm C18 column, 110 Å, 100 x 2 mm. The gradient for all spectra was 3-95% acetonitrile +0.1% formic acid in water.

PCR amplification of *san* genes from pKZ11 and pKZ2

All *san* genes were amplified from the cosmid pKZ11, with the exception of *sanW* which was amplified from cosmid pKZ2.[1] Primers for each gene were designed for use with splicing by overlap extension (SOE) to assemble small contigs. For SOE, each primer must have a 15 base pair or greater overlap with the primer for the gene both to the 3' and 5' end of it in our designed constructs. Additionally, a ribosome binding site was placed in front of each gene. Restriction sites were also incorporated to join larger groups of genes together by standard digestion/ligation methods and to allow facile manipulation of constructs. A complete list of the primers used to amplify all genes can be found in Table S1.

Construction of shuttle vectors

All cloning steps were performed in either *E. coli* XL1-Blue (Stratagene, La Jolla, CA, USA) or *E. coli* DH5α (MCLab, South San Francisco, CA, USA). Plasmids used for cloning included pUC18 (New England Biolabs, Ipswich, MA, USA) and pET28b (Novagen, now EMD Bioscience, Madison,WI, USA). *S. coelicolor* CH999/pBOOST*, which lacks the complete

actinorhodin (*act*) gene cluster, was used as the host for production of polyketides.[2] The plasmid pBOOST* has been shown to co-integrate with vectors containing the SCP2* origin of replication, thereby resulting in higher copy numbers and correspondingly improved antibiotic titers. The transformation of shuttle vectors bearing *san* pathway genes into *S. coelicolor* CH999/pBOOST* was performed following standard procedure.[3]

The shuttle vectors used in this study were constructed in the following manner. Each gene was amplified using the primers listed in Table S1. After amplification of *san*F, *san*G, *and san*H, the 2.8 kb *san*FGH segment was assembled by SOE. For SOE, 50 ng of each gel purified gene in 4 µL water, 0.5 µL of 10 µM *san*F F primer, 0.5 µL of 10 µM *san*H R primer, 4 µL DMSO, 0.5 µL of 20 mM dNTPs, and 10 µl of Phusion GC buffer were added to 30 µL of water. Following this, 0.5 µL of Phusion Hot Start II (Thermo Scientific) was added to the reaction mixture and the reaction was placed in a thermocycler. The thermocycler time program used was as follows: 1. 98°C 3 min; 2. 98°C 10 sec; 3. 60°C 30 sec; 4. 72°C 90 sec; 5. Repeat steps 2-4 30x; 6. 72°C 10 min. The *san*IX, *san*S2WC, and *san*IJ fragments were constructed in a similar fashion. Each of these gel-purified fragments was then digested with an appropriate endonuclease as per Table S1 and ligated into a modified pET28 derived vector, pJF46. The plasmid pJF46 was engineered by replacing the HindIII to EcoRI fragment of pET28b with a small custom oligonucleotide containing restriction sites for PacI, XbaI, SacI, KpnI, and MfeI each separated by six base pairs. The exterior XbaI site was also mutated into a non-cleavable site. The *san*FGH segment was designed with PacI, XbaI restriction sites flanking the genes and was ligated into the corresponding sites of the pRM5 derived shuttle vector pSEK4[4] (which is a derivative of pRM5 with an inactivated *actIII*) to give pJF40. The vector pJF40 was then used as the basis for all additional vectors: pJF77, pJF76, and pJF111. To create these, the additional enzymes for each construct were inserted through digestion of an appropriate JF46-based vector containing the enzymes of interest with XbaI and EcoRI, and ligation into the XbaI/EcoRI site of pJF40.

Production, isolation and characterization of polyketide products from agar plates

Each strain was grown on R5 agar plates containing 50µg/L thiostrepton (Santa Cruz Biotechnologies, Santa Cruz, CA) and 100 µg/L of apramycin (for those strains harboring the pBOOST* plasmid) at 30° C for 7 days. After incubation, plates were extracted 1:1 v:v with EtOAc/MeOH/acetic acid (89:10:1). This crude material was concentrated *in vacuo* and purified

by a variety of methods including preparative HPLC, C_{18} solid phase extraction chromatography, and silica gel chromatography. A representative isolation of TW95b is described here: the dried crude extract (1g) from three liters of R5 agar plates streaked with CH999/pBOOST*/pJF76 was rinsed with 10% acetonitrile in water. The dissolved material was applied to a C18 solid phase extraction cartridge (10 g C18, 60 mL tube, Supelco Discovery) and was washed through with 10-12 column volumes of 10% acetonitrile. The same process was repeated for 33%, 50%, and 100% acetonitrile washes. The 33% acetonitrile extract (350 mg), which contained TW95b, was dissolved in 3 ml of methanol and injected onto a preparative reverse-phase HPLC column (250 mm x 21.2 mm, 5 μm, C18 column, Restek). A gradient of 10-45% acetonitrile + 0.1% TFA in water over 90 min with a flow rate of 5 mL/min was used to elute TW95b. Fractions were collected at 1 min intervals and fractions 69-77 (27-30% acetonitrile, 20 mg) were found to contain TW95b by NMR. These fractions were combined and subjected to semi-preparative HPLC (250 mm x 10 mm, 5 μm, C18 column, Restek) with a gradient of 20-40% acetonitrile +0.1% TFA in water over 45 min with a flow rate of 0.4 mL/min was used to elute TW95b. The fraction from min 22 (30% acetonitrile, 5mg) contained pure TW95b and was used for NMR characterization (Figure S3).

Production, isolation and characterization of polyketide products from liquid cultures

Strains of interest were grown in 1 L R5 media supplemented with 50μg/L thiostrepton in 2 L-baffled flasks. After 4 days of shaking at 30 °C, the supernatant was collected from each culture by centrifugation. Ammonium sulfate was added to a final concentration of 0.5 g/mL, and the solution was stirred at room temperature for 1 h. After further incubation of this solution at 4 °C for 12 h, the precipitate was collected by centrifugation and extracted with MeOH (3 x 4 mL). The MeOH solution was then analyzed by LC-MS for the presence of A-74528. This crude material was concentrated *in vacuo* then purified and analyzed as noted below.

A-74528 analysis by Q-TOF LC-MS

Small scale extracts (~500 mL) were evaporated to dryness and re-suspended in methanol atconcentrations of 10 mg/mL (JF constructs), 25mg/mL (*san* knockout mutants, agar plates) or

50 mg/mL (*san* knockout mutants, ammonium sulfate precipitation). 10 µL of each solution was analyzed by Q-TOF LC-MS. A-74528 elutes at 12.8 min (59% acetonitrile) and displays the molecular ion $[M+H]^+$ *m/z* 561.13927 (calculated for $C_{30}H_{25}O_{11}$, 561.13186). Presence of A-74528 in each extract was evaluated by extracting for ions in the mass range of 561.0-561.2 from the total ion count (TIC). Peaks in the extracted ion chromatogram (EIC) were deduced to be A-74528 based on matching retention time and high resolution molecular formulae, in addition to fragmentation patterns by tandem mass spectrometry. For MS-MS experiments, the ion *m/z* 561.13 with retention times between 12.8 and 12.9 min was selected for fragmentation at 100 V. Diagnostic fragment ions include 375.085, 417.096, 435.1068, and 543.126.

A-74528 isolation from CH999/pBOOST*/pJF111

CH999/pBOOST*/pJF111 was grown on 10 L of R5 agar plates containing 50µg/L thiostrepton (Santa Cruz Biotechnologies, Santa Cruz, CA) and 100 µg/L of apramycin at 30° C for 7 days. After incubation, plates were extracted 1:1 v:v with EtOAc/MeOH/acetic acid (89:10:1). This crude material was concentrated *in vacuo* and the dried crude extract (3.6 g) was dissolved in methanol, adsorbed to loose C_{18} resin (Supelco Discovery DSC-18), and the solvent was then removed. The material was then applied to a C18 solid phase extraction cartridge (10 g C_{18}, 60 mL tube, Supelco Discovery) and was washed through with 10-12 column volumes of 10% acetonitrile. The same process was repeated for 20%, 30%, 35%, 40%, 50% and 100% acetonitrile washes. The 35% and 40% acetonitrile extracts both contained A-74528 and were combined (590 mg). The resulting sample was purified by preparative reverse-phase HPLC (250 mm x 22 mm, 10 µm, C_{18} column, Grace Altima). A gradient of 35-45% acetonitrile + 0.1% TFA in water over 45 min with a flow rate of 12 mL/min was used to elute A-74528. Fractions were collected at 1 min intervals and fractions 25-28 (41% acetonitrile, 31 mg) were found to contain A-74528 by LC-MS. These fractions were combined and subjected to analytical HPLC (250 mm x 4.6 mm, 5 µm, Sync Hydro RP, ESI Industries) with isocratic conditions (37% acetonitrile +0.1% TFA in water over 30 min with a flow rate of 0.8 mL/min), providing semi-pure A-74528 (1.7 mg, 23 min, Figure S7, Table S5).

PCR targeting protocol

Gene disruption in pKZ11 was accomplished using PCR-targeting strategies similar to those previously published.[5] To prepare pKZ11 mutant strains, long PCR primers (58 and 59 nt) were designed such that the 5′ end 39 nt match the pKZ11 sequence adjacent to the gene to be inactivated, and a 3′ sequence (19 or 20 nt) match the right of left end of the disruption cassette, pIJ773, which contains an apramycin resistant gene and oriT site flanked by FLP recombinase sites (Table S3). PCR amplification of the extended resistance cassette was accomplished using the long PCR primers and the purified linear cassette as a template. The amplified extended cassette was transformed into electrocompetent BW25113/λRED/pIJ790 and incubated overnight at 30 °C (pIJ790 contains a temperature sensitive origin of replication) on LB agar plates supplemented with carb (pKZ11) and apramycin (disruption cassette). Plasmids were isolated from a dozen colonies and screened for the insertion of the disruption cassette by PCR amplification with primers pairs just ~100 bp outside the region affected by homologous recombination (see Table S4) and restriction digestion (XhoI). Cosmids containing the extended cassette were transformed into electrocompetent E. coli DH5α/BT340 cells and incubated overnight at 30 C (to retain flipase) on LB agar containing carb (pKZ11), apra (extended cassette), and cml (BT340) for 2 days. Four colonies were restreaked for single colonies on LB/carb and incubated at 42 C to induce the expression of the FLP recombinase followed by the loss of plasmid BT340. Four more colonies per plate were restreaked on LB/carb and incubated at 42 C for single colonies, which were finally plated on two master-plates by streaking 50 single colonies with a toothpick first on LB/carb and then on LB/apra. Colonies that are $carb^R$ and $apra^S$ contain a 81 bp scar in place of the extended resistant cassette, which was further confirmed by PCR amplification of the disrupted region using outside primers and restriction digestion (XhoI).

Mutant cosmids were transformed into *S. lividans* K4114 protoplasts using standard protocols.[3] *S. lividans* K4-114 was selected as the host, because transformation of pKZ11 into *S. coelicolor* CH999 yielded strains that produced no polyketide products. The cosmid pKZ11 contains the entire *san* cluster with the exception of *sanW,* the putative phosphopantatheinyl transferase (PPTase). As transformation of *S. lividans* with pKZ11 did produce polyketide products, we hypothesize that it contains a broad-spectrum PPTase, capable of attaching a phosphopantatheinyl arm onto SanH whereas *S. coelicolor* CH999 does not.

The transformed *S. lividans* protoplasts were plated on R5 agar, and selected by overlaying with thiostrepton (10 μL of the antibiotic in 50 mg/mL antibiotic in 2 mL H_2O) after growing at 30 °C for 18 hours. After 5 days of growth at 30 °C, plasmids were isolated from single colonies using standard protocols[3] and transformed into DH5α. *S. lividans* K4114 colonies harboring the mutant pKZ11 cosmid were confirmed by PCR amplification using the outside primers and restriction digestion (XhoI). Master plates for each mutant strain were obtained upon restreaking positive hits on R5 plates with thiostrepton (50 μg/mL).

ADDITIONAL DISCUSSION

Proposed role of SanJ in A-74528 biosynthesis

To better understand the role of SanJ in the biosynthesis of A-74528, we first examined the function of its closest homologs, members of the Antibiotic Biosynthesis Monooxygenase (ABM) family. Its closest homologs are PdmH (43/58), GrhU (41/59), and BenH (47/58), enzymes involved in the biosynthesis of the pentangular polyphenols pradimycin, griseorhodin, and benastatin, respectively. The role of GhrU in griseorhodin biosynthesis is unknown, but the roles of PdmH and BenH offer clues to the role of SanJ. PdmH acts in concert with the cyclases PdmK and PdmL to affect an oxidation and formation of the C and D rings of pradmicyin[6]. It was found that deletion of either the PdmH oxidase or either of the cyclases abrogates ring formation, yielding the shunt products TW95a and TW95b. BenH has been implicated as a multifunctional oxidase/cyclase containing both an oxidation and a cyclization domain,[7] but its precise role remains unknown.

Based on the role of its homologs, it seems reasonable that SanJ too might serve as an oxygenase and a cyclase in the biosynthesis of A-74528. The putative activities of BenH and PdmH imply that these enzymes act on a full-length polyketide backbone in their respective pathways. This leads us to hypothesize that SanJ also acts after chain formation and possibly also after first ring formation by SanI. Further studies are warranted to determine the precise role of SanJ.

Table S1. Primers used to amplify genes used in pRM5-based plasmids. Primers for each gene were designed for use with splicing by overlap extension (SOE) to assemble small contigs. For SOE, each primer must have a 15 base pair or greater overlap with the primer for the gene both to the 3' and 5' end of it in our designed constructs. Restriction sites are underlined, RBS are in green, start codons are in blue, stop codons are in purple, the 6X His tag on *sanG* is in italics, and mismatches or silent mutations made to allow for easier PCR are in red. The primer used to amplify *sanW* for the minimal PKS construct is denoted with a (77).

Primer	Sequence
sanF F	TTTCCCttaattaaggaggCGGCCGATGACCGGACGCGC
sanF R	ACGCGATCcctcctGCCTTCACGCGATCCTCCTC
sanG F	GTGAAGGCaggaggGATCGCGTG*CATCATCATCATCATCAT*AAGGCGACAAGC
sanG R	ATCTGTTCcctcctGTCACCCGGACCGCACGACCAG
sanH F	CGGGTGACaggaggGAACAGATGAGCAGCAGCATG
sanH R	AGGTGCcctcctctagaATGTCAGACCGCCGTGAGGCT
*sanI*F	TGACATtctagaaggaggACTGACATGGCGCTGACCGACA
sanI R	ATGGGGCCcctcctGAGTCACCGAGCCACCGCCGCGGCACGCTC
sanX F	GGTGACTCaggaggGGACCCATGGCCGTCCGTGGCA
sanX R	ATGGAAGTcctcctgagctcGCGGTTGGTTCAGGCTCCGGGGCAA
*sanS*F	CCAACAGCgagctcaggaggACTTCCATGGAGCCGTCAGCGATTCA
sanS R	CATAGAGCAcctcctACGGACCGGGGTCAGGGCAGGCG
*san2*F	CGGTCCGTaggaggTGGTCTATGACACATCTTCAGCG
san2 R	ATGCGGATcctcctCCCCTATGAGACAGGCAGGTCCAC
sanW F	CATAGGGGaggaggCACAGCATGCGACGCGAAGACGACACCACCACC
sanW R	ATAGTACAcctcctTCAGGACGACGCGACCACCACGGC
sanC F	CGTCCTGAaggaggTGGACGATGCTTCGCCGCGGCCGTGGCA
sanC R	ACCGCGCCcctcctggtaccGCCCCCTCACAGCGCGATGCCGC
sanJ F	GAGGGGGCggtaccaggaggGGCGCGGTGACTCCCGGCCTGCGG
sanJ R	ATCCGGTCcctcctCTCATCCGGTCCTCCGTGACT
sanK F	CGGATGAGaggaggGACCGGATGAGCACGCAGCA
sanK R	AGCACGGCcctcctTCACCCCGTCGTCACACCTCG
sanL F	GGGGTGAaggaggACCGCCATGCCACTGGACGCCA
*sanL*R	ACCCTCGTcctcctAACGTGCTCCGGAACTCAGTCG
sanP F	ATGGCGTTaggaggACGGGGGTGTCTGAGTTGGC
*sanP*R	ATCCGGCCcctcctcaattgTCATCCGGCCCCCTCCACCAGA
sanQ F	CCGGATGAcaattgaggaggGGCCGGATGAGCGCGGCGGCGGGCCGGGT
sanQ R	ATGATCGGcctcctGGTCAGTACGCCGCCGCCACCTC
sanU F	TACTGACCaggaggCCGATCATGTCCACCACTCCTGTCAC
sanU R	TTTAAAgaattcCGCGGCTTCAGGCATGGGTGGGC
san W F(77)	TGACATtctagaaggaggGCACCTATGCGACGCGAAGACGACACCACCA
san W R(77)	GTCAGAcctcttctagaATGTCATCAGGACGACGCGACCACCACGGC

Table S2. All genes in the identified *san* gene cluster are shown arranged by their potential utility in A-74528 biosynthesis.

[a] The required genes are those with homologs in the *fdm* cluster which have been shown to be important for primer unit formation, polyketide backbone formation, or first ring cyclization. [b] The potentially important genes are those which potentially play a role in oxidation of the terminal olefin for A-74528 biosynthesis or may contribute to the unique cyclization pattern of A-74528. [c] Those genes with known homologs in rubromycin-type polyketide biosynthesis (*sanD* and *sanE*) as well as those shown to be active in tailoring reactions in the FDM A pathway most likely will not contribute to overproduction of A-74528. [d] Our laboratory has previously found that overexpression of polyketide products in *S. coelicolor* CH999 using a pRM5-derived vector usually does not require native transporters or regulators.

Gene	Homolog function	Proposed function
Required[a]		
san2	Acyl-ACP thioesterase	purge acetyl-primed ACP
sanC	3-Ketoacyl ACP-reductase	KR for hexadiene formation
sanF	Ketosynthase (KS)	minimal PKS
sanG	Chain Length Factor (CLF)	minimal PKS
sanH	Acyl Carrier Protein (ACP)	minimal PKS
sanI	Polyketide cyclase/aromatase	C9-C14 cyclization
sanS	3-Ketoacyl ACP synthase	hexadiene construction
sanW	Phosphopantetheinyl transferase	holo ACP formation
Potentially important[b]		
sanJ	Monooxygenase	terminal olefin epoxidation
sanK	Monooxygenase	terminal olefin epoxidation
sanL	Monooxygenase	terminal olefin epoxidation
sanP	Monooxygenase	terminal olefin epoxidation
sanQ	Monooxygenase	terminal olefin epoxidation
sanX	Unknown	cyclization/oxygenation
sanU	Unknown	cyclization/oxygenation
FDM specific role[c]		
sanD	Polyketide cyclase	FDM 5th ring cyclization
sanE	Polyketide cyclase	FDM 3rd and 4th ring cyclization
sanM	Hydroxylase	FDM C6 oxidation[8]
sanM1	Hydroxylase	FDM C8 oxidation[8]
sanN	O-methyltransferase	FDM O11 methylation
sanV	Asparagine synthetase	C1 amidation[9]
Unlikely to be required[d]		
san1	Acyl-CoA decarboxylase	Increase malonyl-CoA concentration
san3	Quinone monoxygenase	Not required in FDM A
sanT	Peptide transporter	
sanO	Ketoreductase	Possible C19 KR. Not C3 KR for hexadiene formation
sanT1	Transporter	
sanR	Regulator	
sanR1	Regulator	
sanT2	Transporter	

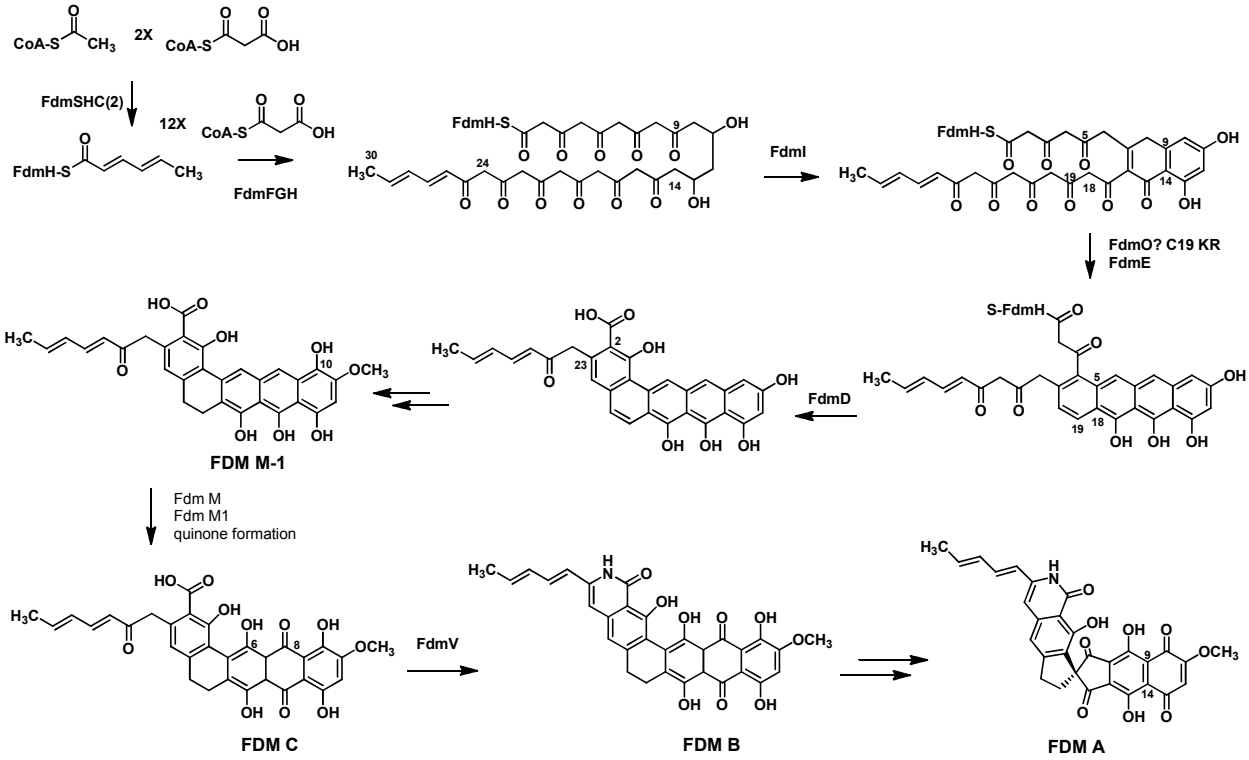
Table S3. Primers used to amplify the extended cassette for PCR-targeting of pKZ11. Long PCR primers (58 and 59 nt) were designed such that the 5′ end 39 nt match the pKZ11 sequence adjacent to the gene to be inactivated (lowercase letters), and a 3′ sequence (19 or 20 nt) match the right of left end of the disruption cassette (capital letters).

Cosmid	Target Gene	Primer	Sequence
pLKC18	*san*3	san3_20_F	ggtggtgacgccgtcgtacgccaggtgcagcaccattccATTCCGGGGATCCGTCGACC
		san3_19_R	atggtggagatcctcttcgaacagaacctgctgtgggccTGTAGGCTGGAGCTGCTTC
pLKC27	*san*J	sanJ_20_F	cccggcctgcgggtgctgctgcggatcgagatcaacgcgATTCCGGGGATCCGTCGACC
		sanJ_19_R	ggtcctccgtgactcaagggttcggccgtcgcggccgggTGTAGGCTGGAGCTGCTTC
pLKC28	*san*K	sanK_20_F	acgcagcaggagcaggtggccgagctgatcgagctggtcATTCCGGGGATCCGTCGACC
		sanK_19_R	tcacacctcggggtggtcgaaccactcgccgagcttggtTGTAGGCTGGAGCTGCTTC
pLKC29	*san*L	sanL_20_F	tcagtcgacacccagtacggacatatgcgtgatcgacagATTCCGGGGATCCGTCGACC
		sanL_19_R	atgccactggacgccatcacctaccgtgtccgacccggcTGTAGGCTGGAGCTGCTTC
pLKC32	*san*O	sanO_20_F	atggaacttcgactggtcggcaagaacgccctggtcaccATTCCGGGGATCCGTCGACC
		sanO_19_R	cccgtcgacgtgcagcgtctcaccgttgacgtacgacgaTGTAGGCTGGAGCTGCTTC
pLKC33	*san*Q	sanQ_20_F	gagcgcggcggcgggccgggtgcggatcgtgctgtatctATTCCGGGGATCCGTCGACC
		sanQ_19_R	tcagtacgccgccgccacctcgaacagcgcgaacggactTGTAGGCTGGAGCTGCTTC
pLKC37	*san*U	sanU_20_F	atgtccaccactcctgtcaccgctcccgccgcgagcccgATTCCGGGGATCCGTCGACC
		sanU_19_R	tcaggcatgggtgggccgttcggcggcctcccgccggaaTGTAGGCTGGAGCTGCTT
pLKC41	*san*P	sanP_20_F	ttggccgggcaggtcgggtccttccgtgtgctgctgaccATTCCGGGGATCCGTCGACC
		sanP_19_R	ggccccctccaccagatgcatcatcgccagcgagccgccTGTAGGCTGGAGCTGCTTC
pLKC48	*san*X	sanX_20_F	gccgtcccccgtgggcgtccctgagccctgggccccatgATTCCGGGGATCCGTCGA
		sanX_19_R	ggtccgtgcgtgagcttccgtggggcggcggttggttcaTGTAGGCTGGAGCTGCTTC

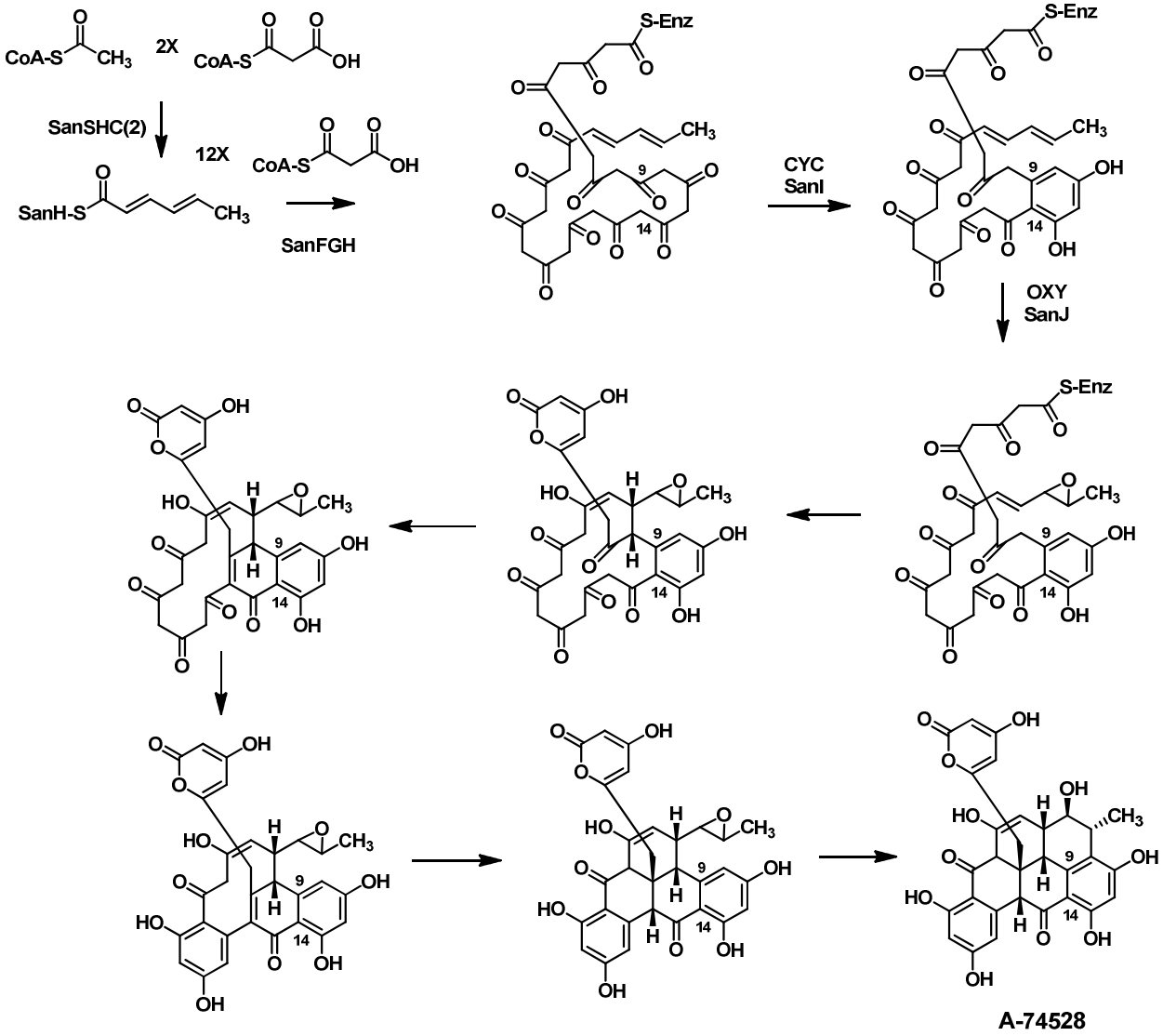
Table S4. Primers used to confirm the integration of the extended cassette and subsequently the 81-bp scar in pKZ11.

				expected amplimer (bp)		
Cosmid	**Target Gene**	**Primer**	**Sequence**	**WT**	**cassette**	**scar**
pLKC18	*san*3	san3_out_F	gggtgctggtcgaggccggtccg	3183	1873	585
		san3_out_R	cagaatccgatgtcgtaccctgc			
pLKC27	*san*J	sanJ_out_F	atgcgctggcggcaggagttcac	640	1754	466
		sanJ_out_R	cttggccttctccgggttgaact			
pLKC28	*san*K	sanK_out_F	gagccgtccttccgcgccttcga	750	1768	480
		sanK_out_R	gctccagaaggccgggttgatcg			
pLKC29	*san*L	sanL_out_F	gctttccgtctcggacttcttct	1190	1896	608
		sanL_out_R	gcaacgtcccgatcaaccgctcg			
pLKC32	*san*O	sanO_out_F	ttcgagctggtgaacgagccgac	1120	1825	537
		sanO_out_R	aacgagtgacggtccgtccagtc			
pLKC33	*san*Q	sanQ_out_F	aggactccgtctactacgtggt	703	1834	546
		sanQ_out_R	atggaggaactgctctcctccgg			
pLKC37	*san*U	sanU_out_F	gcttacgggtcttcgtctgggtc	886	1892	604
		sanU_out_R	cgacgtcgatgaccgccagcctg			
pLKC41	*san*P	sanP_out_F	tatctgcggatgatggcactggg	623	1764	476
		sanP_out_R	cattcgctgagcaccgcgtacgc			
pLKC48	*san*X	sanX_out_F	tcagcgtcgctgtcggccctcga	560	1629	341
		sanX_out_R	ttgatgcggggtcggcttcgccg			

Scheme S1. A proposed biosynthetic pathway to FDM A. Formation of the hexadienyl priming unit is shown from acetyl- and malonyl-CoA for simplicity, though other possibilities exist. [10] The tailoring steps from the earliest isolated intermediate FDM M-1 through FDM A have also been studied by the Shen group. [8]



Scheme S2. A proposed biosynthetic pathway to A-74528. Formation of the hexadienyl priming unit is shown from acetyl- and malonyl-CoA for simplicity, though other possibilities exist.[10] The timing of the oxygenation is purely speculative as it is possible that the terminal olefin could be oxidized as early as during the formation of the priming unit or as late as immediately prior to the final cyclization. The timing of the cyclizations is also speculative and the role of SanJ or other enzyme components in these cyclizations remains unknown. As shown, a series of intramolecular Michael-type additions, are potentially responsible for forming the intricate bonding pattern found in the final product.



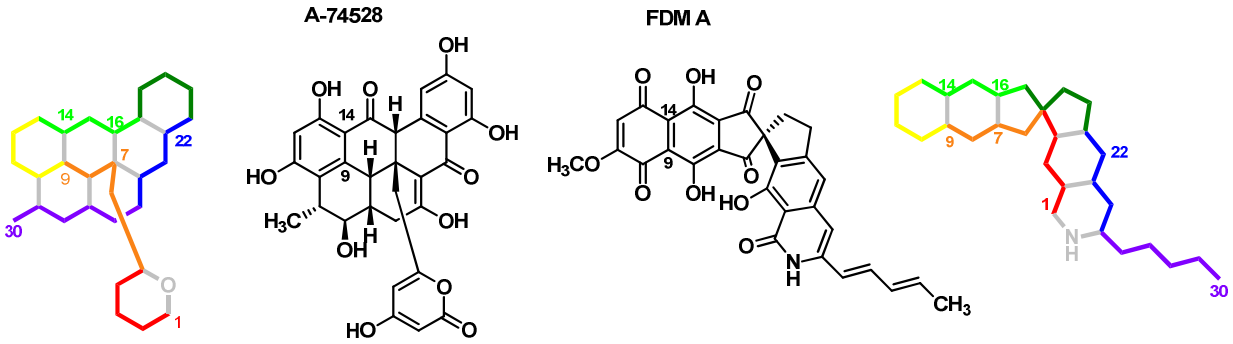


Figure S1. The cyclization patters of A-74528 and FDM A. The two colored representations highlight the difference in cyclization patterns going from C1 (red) to C30 (purple). Formation of a spiro junction in the later stages of biosynthesis of FDM A eliminates C18.

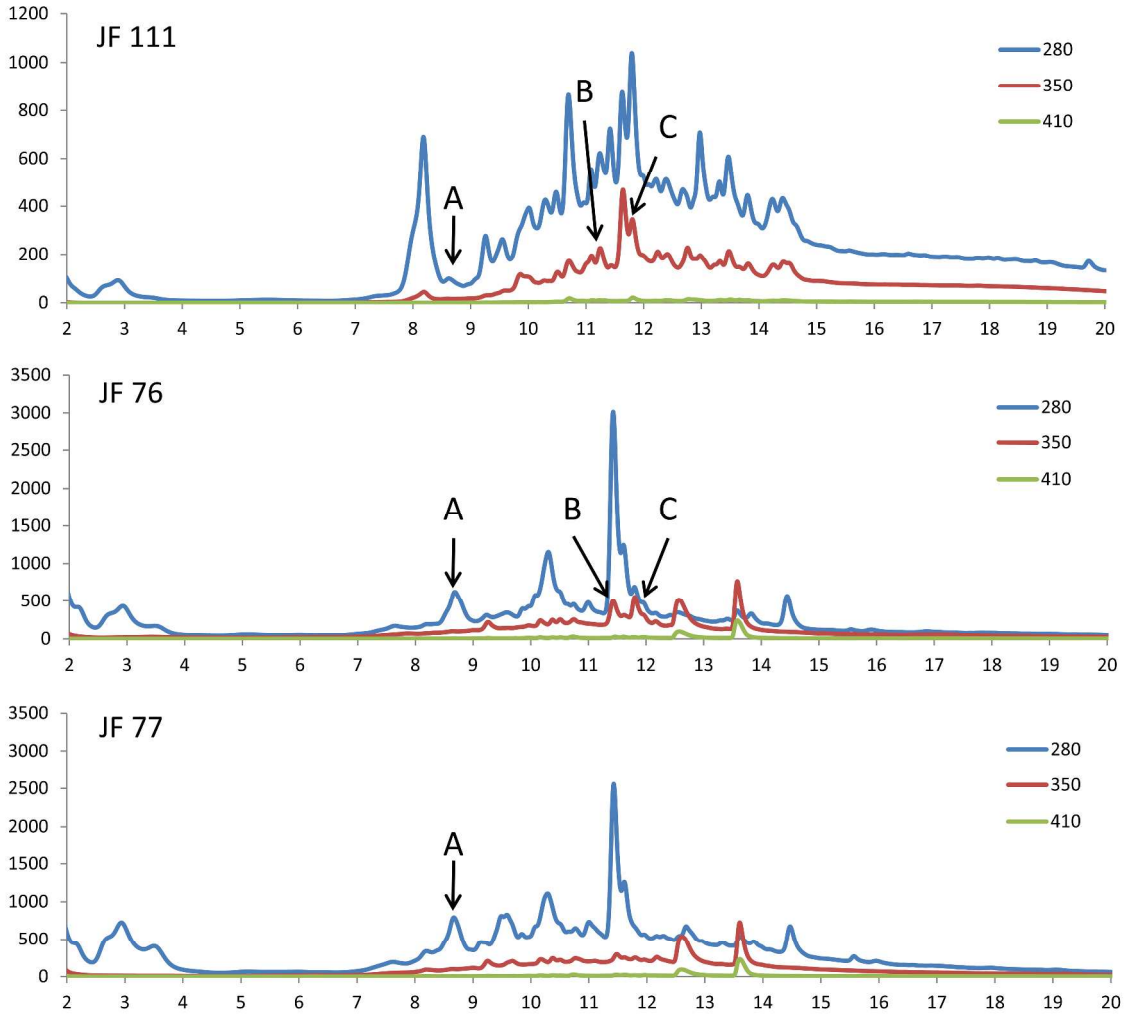


Figure S2. Metabolite profiles of JF77, JF76, and JF111. Peak A in JF77 was isolated and found to contain a triketide and tetraketide by NMR and LC/MS. The m/z for these compounds (127 and 169) was also observed at the same retention times in JF76, and JF111. Peaks B and C, which have absorbance at both 280 and 350 nm, indicating a large conjugated system, appear in JF76, but are not present in JF77. These peaks have retention times consistent with the previously characterized compounds TW95a and TW95b respectively (though in the case of TW95a the peak overlaps with another unidentified product with a large absorbance at 280nm). It is interesting to note that these are the only new peaks to appear in JF76 compared to JF77. This implies that the addition of *sanI* allows enough stabilization of the pocket to begin to form longer-chain products, though they are still not major products. With the addition of *sanJ* in JF111, the TW products are made in isolable quantities (Figure S3).

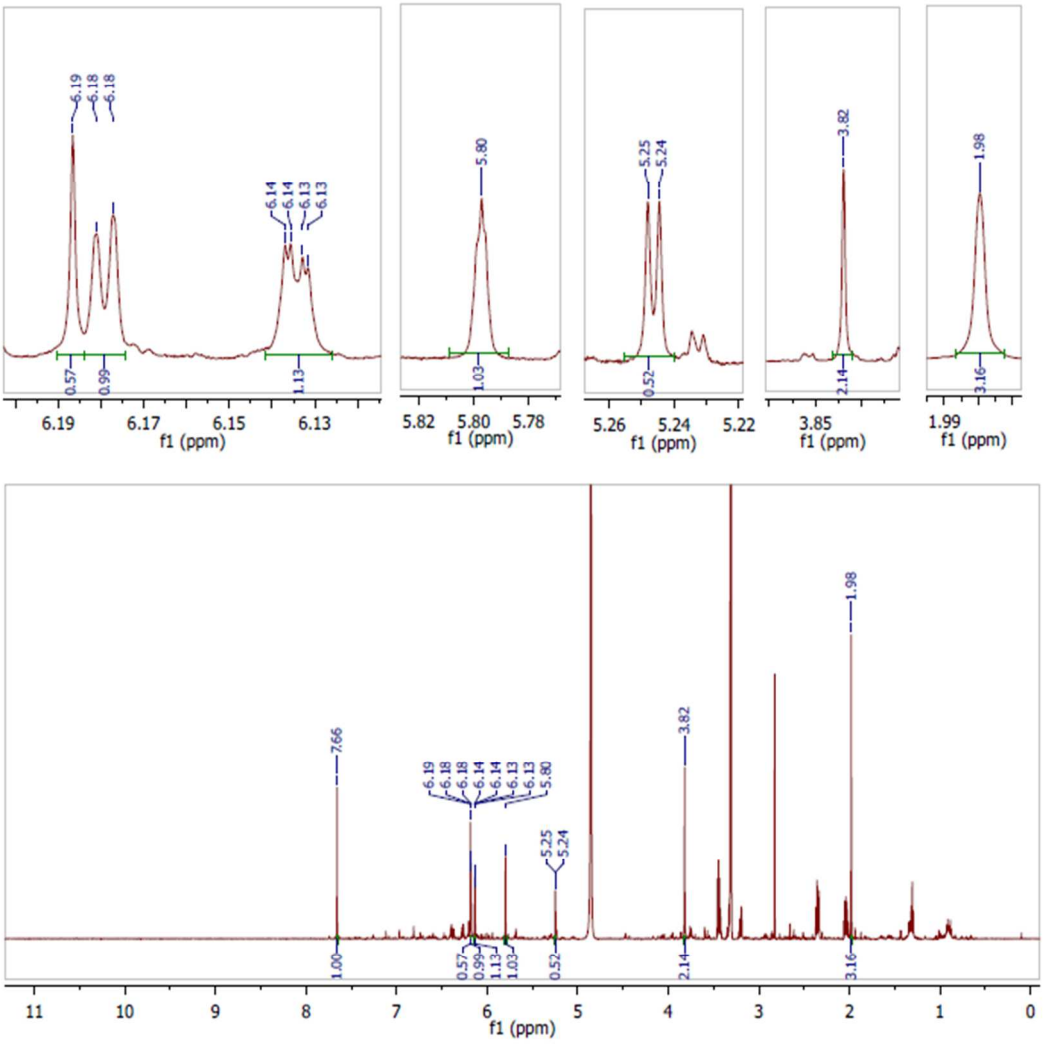


Figure S3. ^{1}H NMR of TW95b. The 600 MHz ^{1}H spectrum of TW95b in CD_3OD is shown above and matches exactly with the literature standard. Enlargements of some key peaks are shown in the top pane and the full spectrum is shown in the bottom panel.

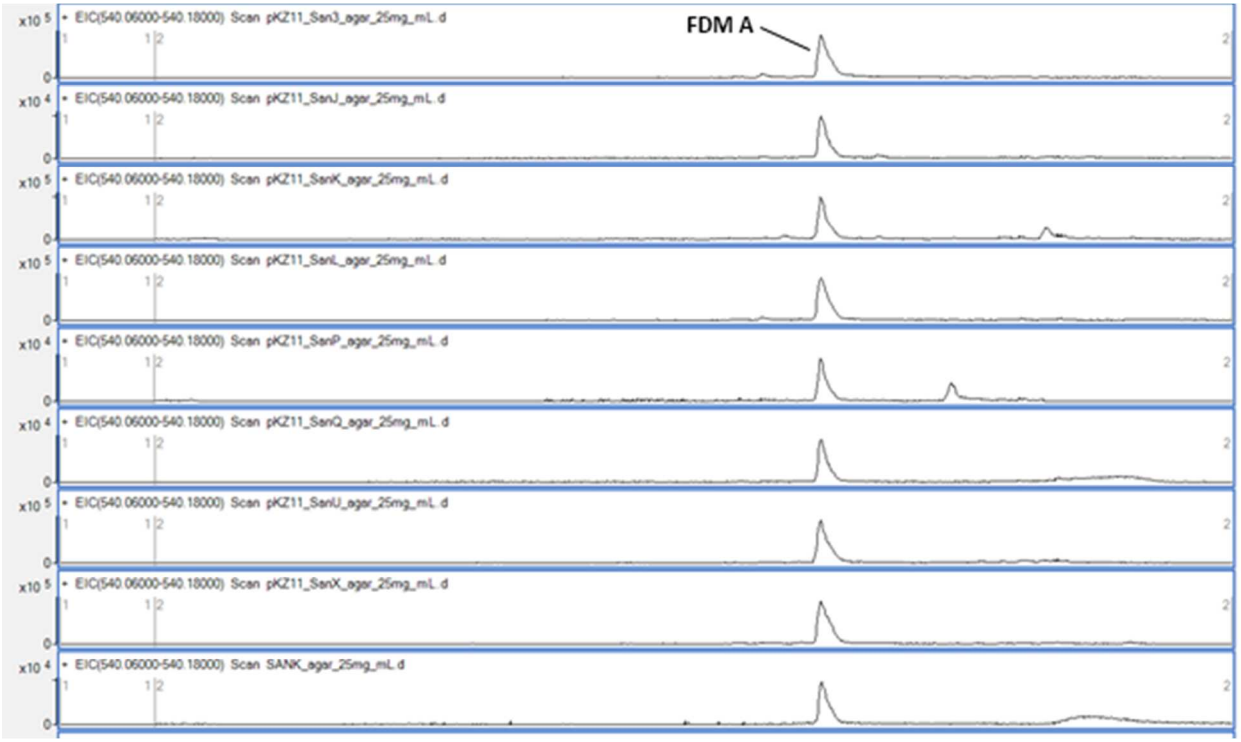


Figure S4. Extracted ion chromatograms for the mass of fredericamycin (FDM A) in Δ*san* mutant strains.

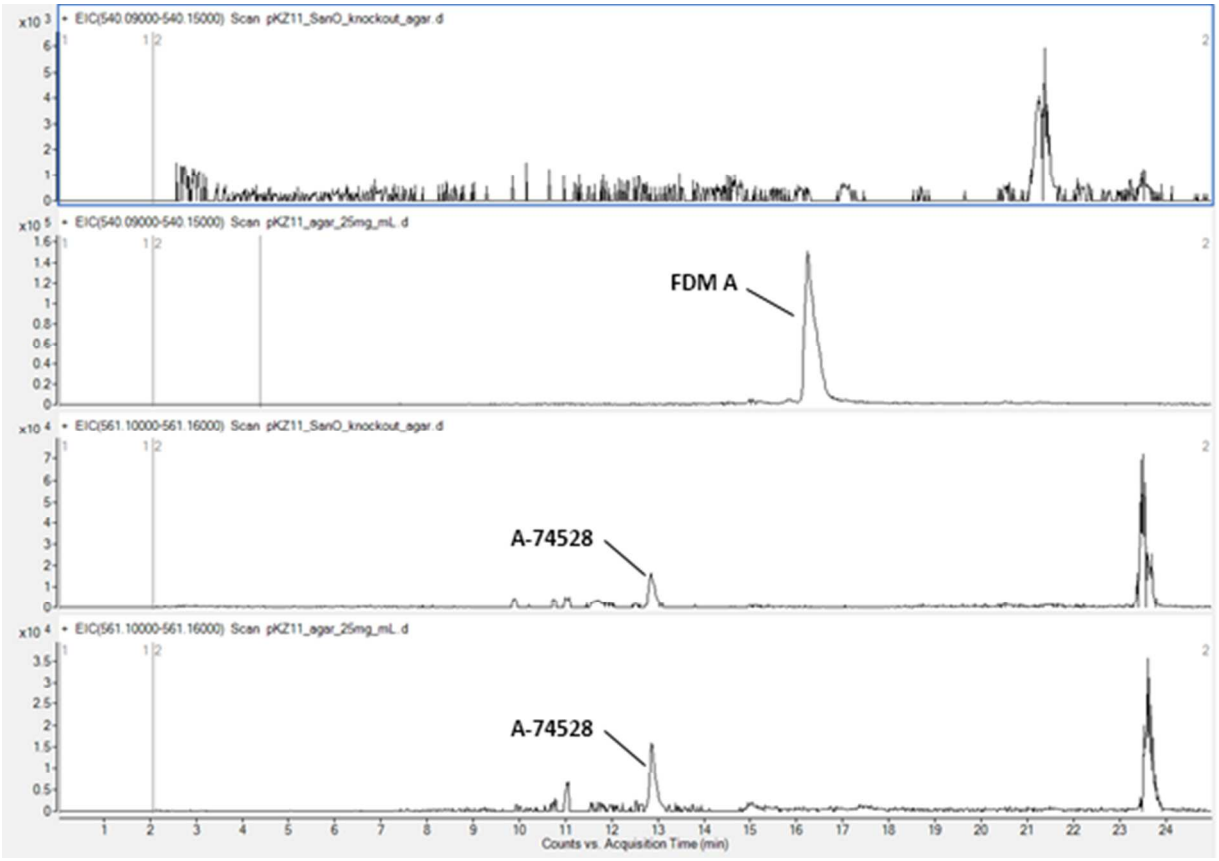


Figure S5. Extracted ion chromatograms (EIC) for A-74528 and fredericamycin (FDM A) in pKZ11 and pKZ11/Δ*sanO*. *S. lividans K4114/*pKZ11 produces both A-74528 and FDM A, whereas *S. lividans K4114/*pKZ11/Δ*sanO* does not produce FDM A.

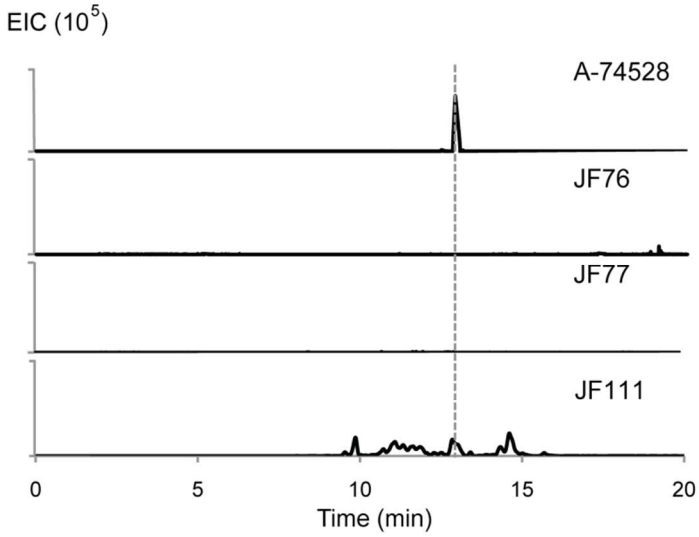


Figure S6. LC/MS analysis of extracts isolated from the heterologous expression strains reveals that CH999/pBOOST*/JF111 harbors the minimal set of genes required for the biosynthesis of A-74528. The peak at 12.85 minutes corresponding to A-74528 in the LC trace was confirmed by MS/MS analysis. The extracted ion chromatogram (EIC) for the positive ion of A-74528 ($[M+H]^+$ 561.1-561.15) is shown on the Y-axis.

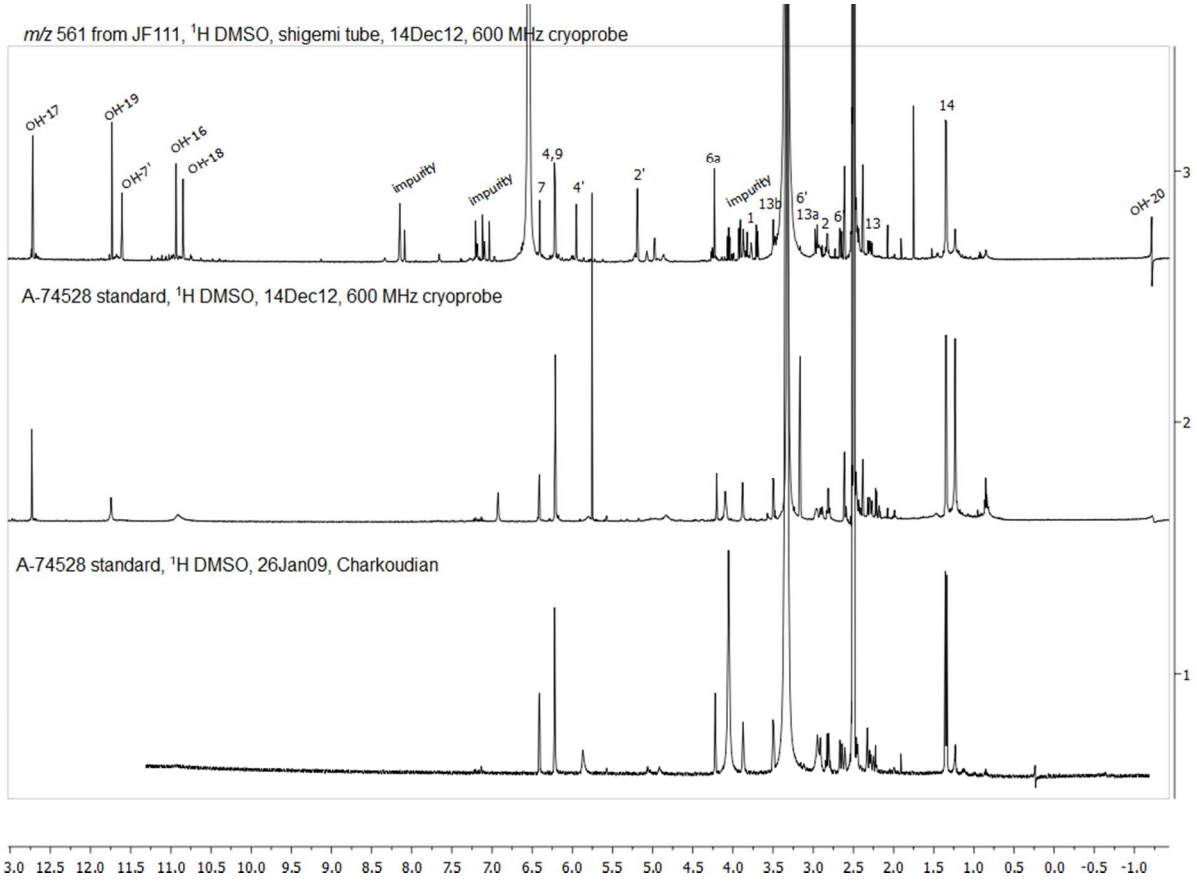
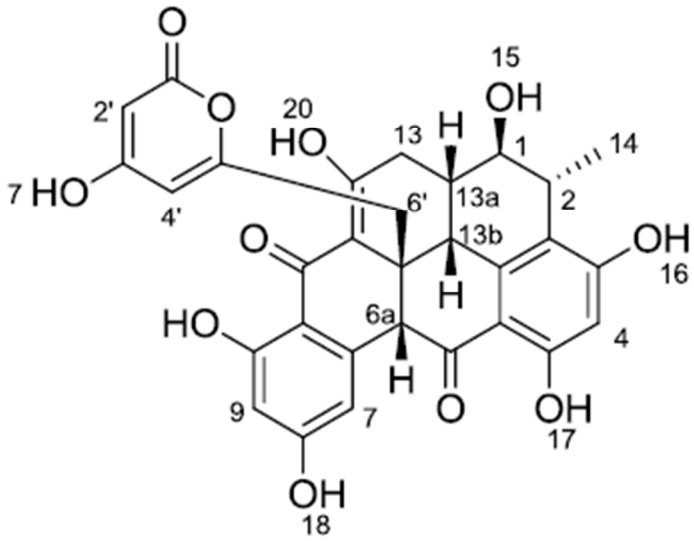


Figure S7. Proton NMR of A-74528 comparison to standard. The ^{1}H NMR of A-74528 isolated from JF111 shows good agreement with the standard as well as literature values. The isolated compound from JF111 is shown in the top panel and compared two spectra of our authentic standard in the lower panels. Spectra is recorded in a 5 mm shigemi tube matched to DMSO-d_6 on a Varian Inova 600 MHz NMR, outfitted with a cryoprobe (University of California, Santa Cruz).



	Literature		A-74528 Standard			m/z 561,isolated from JF111		
Proton #	δ_H	J (Hz)	δ_H	J (Hz)	ΔLiterature	δ_H	J (Hz)	ΔLiterature
1	3.87	d, 3.2	3.88	m	0.01	3.87	m	0.00
2	2.82	q, 7.5	2.82	q, 7.4	0.00	2.82	q, 7.5	0.00
OH-16	10.90	br s	10.92	br s	0.02	10.93	s	0.03
4	6.22	s	6.21	s	0.01	6.22	s	0.00
OH-17	12.71	s	12.73	s	0.02	12.72	s	0.01
6a	4.23	s	4.20	s	0.03	4.23	s	0.00
7	6.40	d, 2.2	6.41	d, 1.9	0.01	6.40	d, 2.3	0.00
OH-18	10.82	br s	not visible			10.85	s	0.03
9	6.20	d, 2.2	6.21	d, 2.2	0.01	6.21	d, 2.3	0.01
OH-19	11.73	br s	11.74	s	0.01	11.73	s	0.00
OH-20	14.80	s	14.78	br s	0.02	14.79	s	0.01
13	2.33	dd, 19.1, 12.0	2.30	dd, 19.3, 12.1	0.03	2.30	dd, 19.1, 12.1	0.03
	2.49	dd, 19.1, 6.1	signal overlap with DMSO			signal overlap with DMSO		
13a	2.94	m	2.96	m	0.02	2.92	m	0.02
13b	3.50	d, 3.4	3.50	d, 3.8	0.00	3.50	d, 3.9	0.00
14	1.34	d, 7.5	1.35	d, 7.5	0.01	1.35	d, 7.5	0.01
2'	5.18	d, 1.8	5.06	br s	0.12	5.19	d, 2.1	0.01
OH-7'	11.56	br s	not visible			11.61	s	0.05
4'	5.94	d, 1.8	5.87*	br s	0.07	5.95	d, 2.1	0.01
6'	2.65	d, 14.7	2.60	d, 14.2	0.05	2.66	d, 14.7	0.01
	2.98	d, 14.7	2.90	d, 14.2	0.08	2.97	d, 14.7	0.01
			Average ΔLiterature		0.03			0.01

Table S5. H^1 NMR comparison of A-74528. The A-74528 isolated from JF111 shows very good agreement with both the literature values[11] and the NMR of our authentic standard.

Supplemental Information References

(1) Zaleta-Rivera, K.; Charkoudian, L. K.; Ridley, C. P.; Khosla, C. *J Am Chem Soc* **2010**, *132*, 9122–8.

(2) McDaniel, R.; Ebert-Khosla, S.; Hopwood, D.; Khosla, C. *Science* **1993**, *262*, 1546–1550.

(3) Kieser, T., Bibb, M.J., Buttner, M.J., Chater, K.F., and Hopwood, D. A. *Practical Streptomyces Genetics*; The John Innes Foundation: Norwich, UK, 2000; pp. 472–481.

(4) Fu, H.; Ebert-Khosla, S.; Hopwood, D. A.; Khosla, C. *J Am Chem Soc* **1994**, *116*, 4166–4170.

(5) Gust, B.; Kieser, T.; Chater, K. F. *REDIRECT Technology. PCR Targeting Systems in Streptomyces coelicolor*; John Innes Centre: Norwich Research Park: Colney, Norwich NR4 7UH, U.K, 2002.

(6) Zhan, J.; Watanabe, K.; Tang, Y. *Chembiochem* **2008**, *9*, 1710–5.

(7) Xu, Z.; Schenk, A.; Hertweck, C. *J Am Chem Soc* **2007**, 6022–6030.

(8) Chen, Y.; Wendt-Pienkoski, E.; Rajski, S. R.; Shen, B. *J Biol Chem* **2009**, *284*, 24735.

(9) Chen, Y.; Wendt-Pienkowski, E.; Ju, J.; Lin, S.; Rajski, S. R.; Shen, B. *J Biol Chem* **2010**, *285*, 38853–60.

(10) Das, A.; Szu, P.-H.; Fitzgerald, J. T.; Khosla, C. *J Am Chem Soc* **2010**, *132*, 8831–3.

(11) Fujita, Y.; Kasuya, A.; Matsushita, Y.; Suga, M.; Kizuka, M.; Iijima, Y.; Ogita, T.; *Bioorg Med Chem Lett* **2005**, 15, 4317-21.